Review Article

Diagnostic accuracy of preoperative prediction of malignancy in ovarian cysts

Introduction

In this review, we will critically evaluate the evolution of the diagnostic tools used in the prediction of ovarian malignancy preoperatively, that has shown to get a better outcome for ovarian cancer patients. It will triage women to the best place for their surgery and by ensuring the expertise available to perform the surgery reducing the rate of incomplete primary surgery, that has reflected into improve survival of ovarian cancer patients.

Robert Jones,[1] Radwan Faraj[2]
[1]College of Medicine, University of Sheffield, UK
[2]Rotherham General Hospital, UK

Correspondence: Radwan Faraj, Rotherham General Hospital, Rotherham, South Yorkshire, United Kingdom, Tel 0044-1709-424191, Email faraj68@hotmail.com

The Risk of Malignancy Index (RMI) was the first diagnostic tool developed to help clinicians differentiate between benign and malignant ovarian tumors. Jacob et al.[1] described how it was developed and assesses its performance amongst a cohort 143 patients referred to a London hospital for elective surgical investigation of an adnexal mass. Each patient had five criteria assessed preoperatively: age, menopausal status, and clinical impression score (as determined by an experienced gynecologist using all of the patient's preoperative clinical information; out of 5, high=more likely to be malignant), ultrasound score (out of 5, high=more likely to be malignant), and CA-125 level.

Following surgery, it was determined that 101 of the patients had benign masses and 42 had malignant tumours. Patients in the malignant group had a higher mean age, were more commonly postmenopausal, and had higher clinical impression scores, ultrasound scores and CA-125 levels than patients in the benign group; all of these differences were statistically significant. However, when used individually, the criteria had inadequate sensitivity and specificity for malignancy.

Logistic regression analysis revealed that only menopausal status, ultrasound score and CA-125 levels were significantly and independently related to the likelihood ratio for malignancy. These three criteria were then combined to form the RMI:

RMI=U×M×CA-125

U=0 for ultrasound score of 0

=1 for ultrasound score of 1

=3 for ultrasound score of 2-5

M=1 if premenopausal

=3 if postmenopausal

When a cutoff score of 200 was used, the RMI showed excellent diagnostic performance, with a sensitivity of 85% and a specificity of 97%. Patients with an RMI score in excess of 200 had 42 times the background risk of cancer, whilst those with a score less than 200 had 0.15 times the risk. The authors concluded that the RMI was a valuable tool for the preoperative assessment of benign and malignant adnexal masses, and that it gave a higher level of discrimination than could be achieved by individual criteria alone.

Overall, this is a good study. The sample size was sufficiently large to allow statistically significant differences between groups to be seen. The criteria used for the RMI were not chosen at random, or with clinical intuition alone: the formula was derived using objective statistical methods (which are presented as an appendix of the manuscript).
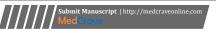
However, the RMI does not take into account any symptoms or signs of ovarian cancer: the clinical impression score was not found to be both statistically and independently related to the likelihood ratio for ovarian cancer, and so was disregarded from further calculations. However, the score was determined by only one gynecologist and the method of calculation is not given, suggesting a high level of subjectivity. More thorough assessment of clinical features–perhaps using a checklist system–might have shown a stronger correlation with malignancy.

Unlike menopausal status and CA-125 levels, the ultrasound score is subjective: it relies on the expertise of the examiner. However, scores were determined by more than one operator in the hospital's radiology department. The article does not consider whether there could have been variability between the different examiners when it came to determining the scores. The International Ovarian Tumour Analysis (IOTA) group have developed a set of simple ultrasound rules, which can be used to preoperatively determine benignity/malignancy in adnexal masses, as an alternative to the RMI.[2] The ultrasonic features to predict a malignant tumour (M features) are: irregular solid tumour (M1), ascites (M2), at least four papillary structures (M3), irregular multilocular solid tumour with a largest diameter of at least 100 mm (M4), and very high colour content on colour Doppler examination (M5).

The five ultrasonic features to predict a benign tumour (B features) are: unilocular cyst (B1), presence of solid components for which the largest solid component is <7mm in largest diameter (B2), acoustic shadows (B3), smooth multilocular tumour (B4), and no detectable blood flow on Doppler examination (B5).

Accordingly, the rules were used as follows:

a. **Rule 1:** if one or more M features are present in absence of B feature, mass is classified as malignant.

b. **Rule 2:** if one or more B features are present in absence of M feature, mass is classified as benign.

c. **Rule 3:** if both B and M features are present, or if no B or M

features are present, result is inconclusive and second stage test is recommended.

(Second stage tests were the RMI, two logistic regression models, and subjective assessment of ultrasonic findings by an experienced ultrasound examiner.)

To assess the diagnostic performance of their simple ultrasound rules, the IOTA group carried out a prospective temporal and external validation study amongst 1,938 patients with an adnexal mass. The results discussed below are those of premenopausal patients only, although postmenopausal patients were entered in the study as well.

To assess the diagnostic performance of their simple ultrasound rules, the IOTA group carried out a prospective temporal and external validation study amongst 1,938 patients with an adnexal mass. The results discussed below are those of premenopausal patients only, although postmenopausal patients were entered in the study as well.

The simple rules yielded a conclusive result (either benign or malignant) in 969 (81%) of the 1,196 premenopausal patients who entered the study. In this group, they had a sensitivity of 91% and a specificity of 97%.

Correspondingly, the rules yielded an inconclusive result in 227 (19%) of the premenopausal patients. These patients' masses were then put through second stage tests in accordance with Rule 3: of these, the subjective assessment had the highest sensitivity and specificity.

In summary, the study showed that the simple ultrasound rules could correctly characterise the majority of premenopausal adnexal masses as being either benign or malignant. The simple rules can thus be used as a triage test, and subsequent assessment by an experienced ultrasound examiner can be used as a second stage test in those masses for which the simple rules yield an inconclusive result–if used in this way, the diagnostic performance is very high and there is a reduced number of masses which need referral for expert scanning.

The study's major strength was that it was conducted in 19 ultrasound centres in eight different countries. This multicentre design means that the results were derived from patients with different characteristics, and thus are likely to be applicable in many different settings. The main drawback of the study was that all the examinations were done by experienced ultrasound examiners. Since the rules are supposed to be simple enough to be used by all sonographers and general gynaecologists, their validation by less experienced examiners is needed.

A more recent study from Brazil,[3] the authors questioned whether the presence or absence of certain ovarian cancer symptoms could reliably distinguish between malignant and benign ovarian masses. 176 women with adnexal masses attending the Obstetrics and Gynaecology department of a Brazilian hospital filled out a (previously validated) symptom questionnaire, and had blood levels of CA-125 and HE4 (another ovarian cancer biomarker) measured. Importantly, this was all done preoperatively, meaning that recall bias was avoided.

Following surgery, 60 of the women were diagnosed with ovarian cancer and 116 with benign ovarian masses. As could be expected, the patients with ovarian cancer showed a higher frequency of symptoms– such as pelvic pain, early satiety and abdominal bloating–than those with benign masses.

Following this, a set of six symptom clusters were formed. The three clusters with the highest sensitivity for ovarian cancer were as follows:

A. Abdomen (abdominal bloating and/or increased abdominal size)

B. Pain (pelvic and/or abdominal pain)

C. Eating (unable to eat normally and/or early satiety)

These were then combined to form a symptom index: that is, presence of at least one of the symptoms included in the three clusters. The symptom index showed good performance for differentiating malignancy from benignity (78% of women with ovarian cancer had a positive symptom index, and 60% of women with benign masses had a negative symptom index), and this performance was further enhanced when CA-125 levels were taken into account. HE4 levels were less informative than those of CA-125.

This study should be commended for its lack of recall bias, and for the fact that it has attempted to see whether malignant and benign ovarian masses can be differentiated using simple, readily-available methods. Major drawbacks are that premenopausal and postmenopausal women were not analysed separately, and that this was a single-centre study which analysed a relatively small number of patients.

Another study, carried out across four UK hospitals, aimed to identify diagnostic factors for ovarian cancer.[4] Interestingly, it used both qualitative and quantitative methods. 124 women referred to hospital with suspected ovarian malignancies were interviewed either prior to or shortly after diagnosis. A thematic analysis of the interview data was conducted, and symptoms were quantitatively analysed to assess their sensitivities and specificities for ovarian cancer.

Final diagnoses comprised 44 malignancies and 80 non-cases. Several symptoms–including persistent abdominal distension, appetite loss and early satiety–were reported more frequently in the ovarian cancer group, and these differences were statistically significant. The qualitative aspect to this study was excellent, for two reasons. Firstly, it enabled a deeper understanding of symptom experiences. All of the women diagnosed with cancer reported symptoms prior to diagnosis, but most women saw them as normal changes attributable to ageing or weight gain. This goes some way to explaining why many women present to medical services with advanced ovarian cancer, and helps to debunk the theory that ovarian cancer is a "silent killer".

Secondly, the study showed that the term "bloating", frequently used by women, represented two discrete events: persistent abdominal distension and fluctuating abdominal distension/discomfort. Whilst persistent distension, in the presence or absence of fluctuating distension/discomfort, was significantly associated with ovarian cancer (38 out of the 44 patients reported it), fluctuating distension/ discomfort in the absence of persistent distension was not. This is important, because bloating is frequently referred to as a symptom of ovarian cancer. Previous studies have not made the distinction between persistent and fluctuating distension, and may have therefore misinterpreted symptom information from patients. Furthermore, this finding has implications for GPs–it suggests that they need to fully characterise what a woman means by bloating before assessing her risk for ovarian cancer.

A limitation of this study was that many participants were interviewed after diagnosis, therefore introducing an element of recall bias–women diagnosed with cancer might have over-reported symptoms in retrospect, and vice-versa for those diagnosed as cancer-free. Furthermore, the study had a rather small sample size, mainly because of the time-consuming nature of the interviews. Finally, some of the 80 women without cancer had a benign adnexal mass, but this

group was not analysed separately. Had this not been the case, useful comparisons could have been made between patients with ovarian cancer, patients with benign adnexal masses, and patients without any pathology.

So far, we have discussed methods by which benign and malignant adnexal masses can be differentiated pre-operatively. However, some patients will undergo laparoscopic surgery for an adnexal mass that is expected to be benign, only for the mass to be intra- or post-operatively determined to be malignant. This scenario is known as unexpected ovarian malignancy.

A recent study carried out in a Japanese teaching hospital, aimed to determine the frequency of unexpected ovarian malignancy and whether it affects prognosis.[5] It involved a retrospective review of the pathology reports of 884 patients who underwent laparoscopic resection of adnexal masses. These masses were assessed pre-operatively with ultrasound and were expected to be benign. Cases converted to laparotomy due to pelvic adhesions were excluded from analysis.

Final pathology reports indicated unexpected ovarian malignancy in 13 of the 884 patients (1.5%). 10 of these patients were premenopausal, and 7 of them went on to have a laparotomy at a later date. All patients were alive and without evidence of disease at the last follow-up, the mean duration of which was 38 months (range: 7–80 months).

The authors concluded that unexpected ovarian malignancy was rare amongst their cohort of patients, who were pre-operatively determined as being low-risk for malignancy. As mentioned previously, patients should generally not undergo laparoscopic surgery unless their risk of malignancy is low. Therefore, this study was carried out amongst an appropriate cohort, and the observed rate of unexpected ovarian malignancy seems accurate. Had the pre-operative assessment of masses been less stringent, one would expect the rate to have been higher.

This study is limited by the fact that cases converted to laparotomy were excluded from analysis, and the number of such cases is not given in the manuscript. Some of these patients might have had unexpected ovarian malignancy, and their inclusion in analysis might have given a more representative rate. In addition, the authors observed that the presence of unexpected ovarian malignancy did not alter their patients' prognoses. However, there were only 13 unexpected ovarian malignancies and they were all early stage (1A-C), making estimations of prognosis beyond this study difficult.

To conclude, it seems there are no 'best' screening tools yet for predicting malignancy in ovarian masses. We believe the Multimodel approach that calculates the risk malignancy index using CA125 and ultrasound features will provide a practical and relatively 'safe' way for triaging patients by primary care doctors. Further randomised trials are needed to establish the best screening multimodel.

Synthesis of the literature

Pre-operative estimation of malignancy for adnexal masses is important, as it allows appropriate surgical management to be determined. Laparoscopic surgery is generally not advised for suspected ovarian malignancy: it does not allow for full exploration of the abdominal cavity, and is associated with intra-operative rupture of cyst contents and worsening of prognosis.[6]

In 2011, the Royal College of Obstetricians and Gynecologists (RCOG) published a Green-top Guideline on the management of suspected ovarian masses in premenopausal women.[7] In this, they state that "*an estimation of the risk of malignancy is essential in the assessment of an ovarian mass*". They recommend that the RMI is used for this purpose, with the IOTA simple ultrasound rules as an alternative. Given that the simple rules display greater sensitivity and specificity than the RMI2, it would not be surprising to see a shift in practice in the future.

Unfortunately, there are few studies in this area focusing specifically on premenopausal women, a fact that is highlighted in the RCOG guidelines.[7] CA-125 is possibly a poor predictor of malignancy in this patient group, because of its relatively low specificity (it may be elevated in many other conditions). A method not involving CA-125 might be superior; indeed, the IOTA simple rules have both sensitivity and specificity in excess of 90% amongst premenopausal women.[2]

Acknowledgments

None.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

1. Jacobs I, Oram D, Fairbanks J, et al. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *Br J Obstet Gynaecol*. 1990;97(10):922–929.

2. Timmerman D, Ameye L, Fischerova D, et al. Simple ultrasound rules to distinguish between benign and malignant adnexal masses before surgery: prospective validation by IOTA group. *BMJ*. 2010;341:c6839.

3. Pitta Dda R, Sarian LO, Barreta A, et al. Symptoms, CA125 and HE4 for the preoperative prediction of ovarian malignancy in Brazilian women with ovarian masses. *BMC Cancer*. 2013;13:423.

4. Bankhead CR, Collins C, Stokes-Lampard H, et al. Identifying symptoms of ovarian cancer: a qualitative and quantitative study. *BJOG*. 2008;115(8):1008–1014.

5. Matsushita H, Watanabe K, Yokoi T, et al. Unexpected ovarian malignancy following laparoscopic excision of adnexal masses. *Hum Reprod*. 2014;29(9):1912–1917

6. Schorge JO, Eisenhauer EE, Chi DS. Current surgical management of ovarian cancer. *Hematol Oncol Clin North Am*. 2012;26(1):93–109.

7. RCOG. Management of suspected ovarian masses in premenopausal women. Green-top Guideline No. 62. London; 2011.